

学校编码: 10384

分类号\_\_\_\_\_密级\_\_\_\_\_

学号: X2010230021

UDC\_\_\_\_\_

厦门大学

工 程 硕 士 学 位 论 文

基于 Lucene 的教学资源管理系统  
设计与实现

Design and Implementation of Teaching Resources  
Management System Based on Lucene

罗斌

指导教师姓名: 曾 文 华 教 授

专 业 名 称: 软 件 工 程

论文提交日期: 2013 年 4 月

论文答辩日期: 2013 年 5 月

学位授予日期: 2013 年 6 月

指 导 教 师: \_\_\_\_\_

答辩委员会主席: \_\_\_\_\_

2013 年 4 月

## 厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为( )课题(组)的研究成果,获得( )课题(组)经费或实验室的资助,在( )实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

## 厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

☐ 1.经厦门大学保密委员会审查核定的保密学位论文，于 年 月 日解密，解密后适用上述授权。

☒ 2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

## 摘 要

随着教学信息化的发展,教师授课的方式有了很大的改变,原来的纸质手写教案都被电子版的 PPT 文档所代替。电子版的教学课件方便了学生的下载阅读,能够更好的进行课前预习和课后复习等自学活动。但是随着电子课件的逐年积累,存量越来越大,原有的基于 FTP 服务器的文档管理模式已经不能很好的满足管理和使用的需求。当需要在众多的文档中寻找需要的资料时,只能逐个目录翻查,通过标题猜测文档内容的时候,对文档进行全文检索和在线阅读的需求就显得很迫切。本系统就是为了满足此类需求而产生的。

本文针对教学资源全文检索及在线展示等问题,基于目前已经比较成熟的开源项目 Lucene、Solr、OpenOffice、FlexPaper 等,设计并实现了教学资源管理系统,其主要研究内容如下:

1、本文基于 J2EE 三层架构和 Mysql 数据库设计并实现了教学资源管理系统,主要包括文档转换、文档查看、文档检索、文档栏目管理和用户权限管理等功能模块,重点解决了文档分类管理、全文检索及在线阅读等关键性需求。

2、以软件工程的开发方法为设计主线,较为详细的介绍了教学资源管理系统的业务需求、用户需求、功能性需求、非功能需求、系统架构设计、功能设计和数据库设计。并针对文档转换、文档查看、文档检索等主要功能给出了系统的实现细节及关键功能模块代码,最后给出了系统的相关测试结果。

经过本项目的研发实施将为学院提升教学资源管理效率建立基础信息化平台,较为明显的改变目前资源利用率低下的状况,为广大师生的教学科研提供助力。

**关键词:** 教学资源; 全文检索; 在线展示; Lucene

## Abstract

With the development of education informationization, has undergone great changes in the way teachers teach, the original paper handwriting lesson plans have been replaced by the PPT document. The electronic version of the courseware students download and read helps them in self-learning activities, after-school review. However, with more and more electronic courseware being created, FTP server-based document management mode can not meet the needs of management and use. When finding the right documents from a lot of documents becomes a game of guessing, the urgent demand for full-text search and reading online is obvious. The present system is just to the to meet such demand.

This paper focus on the usage of full-text search of teaching resources and online presence, based on Lucene, Solr, OpenOffice, FlexPaper, etc. It design and implement the teaching resource management system, which main researching contents are as follows:

This paper based on the J2EE three-tier architecture and the Mysql database design and implement of the teaching resource management system, which mainly including the document conversion, document viewing, document retrieval, the document columns management and user rights management and other functional modules. It focuses on solving the the document classified management, the full-text search and online reading.

This paper describes the business requirements, user requirements, functional requirements, non-functional requirements, system architecture design, functional design and database design. It highlights the main functions of document conversion, document viewing, document retrieval. It also shows the tests and results of the system.

After the research and development of the project will contribute to college to upgrade the teaching efficiency in resource management, the more obvious changes be

seen in resource utilization.It will give much help for teachers and students learning and researching.

**Key words:** Teaching resource; Full-text search; Online reading; Lucene

厦门大学博硕士论文摘要库

# 目 录

<b>第一章 引言 .....</b>	<b>1</b>
1.1 研究背景和意义 .....	1
1.2 国内外研究现状 .....	1
1.3 论文的主要研究内容 .....	2
1.4 论文结构安排 .....	2
<b>第二章 相关技术介绍 .....</b>	<b>4</b>
2.1 J2EE技术平台 .....	4
2.2 SSH2 框架 .....	4
2.3 Struts2 概述 .....	5
2.4 Spring概述 .....	6
2.5 Hibernate概述 .....	7
2.6 Lucene全文检索引擎 .....	8
2.7 IK Analyzer分词工具包 .....	11
2.8 Apache Solr服务器 .....	11
2.9 Nginx服务器 .....	12
2.10 Mysql数据库 .....	13
2.11 本章小结 .....	14
<b>第三章 系统需求分析 .....</b>	<b>15</b>
3.1 系统业务需求分析 .....	15
3.2 系统用户需求分析 .....	15
3.3 系统功能性需求分析 .....	17
3.4 系统非功能性需求分析 .....	19
3.5 本章小结 .....	19
<b>第四章 系统总体设计 .....</b>	<b>20</b>
4.1 系统架构设计 .....	20

4.2	系统功能模块设计 .....	23
4.3	系统数据库设计 .....	23
4.4	本章小结 .....	27
<b>第五章</b>	<b>系统详细设计与实现 .....</b>	<b>28</b>
5.1	系统开发与运行环境 .....	28
5.1.1	系统开发环境 .....	28
5.1.2	系统运行环境 .....	28
5.2	文档转换模块 .....	28
5.3	文档查看模块 .....	30
5.4	文档检索模块 .....	32
5.4.1	文档各属性字段的检索 .....	32
5.4.2	全文检索的基本原理 .....	32
5.4.3	全文检索引擎的选择与配置 .....	35
5.4.4	全文检索功能的实现 .....	36
5.5	用户权限管理模块 .....	38
5.5.1	用户登录验证 .....	38
5.5.2	用户密码找回 .....	40
5.5.3	系统角色管理 .....	41
5.5.4	系统用户管理 .....	41
5.6	文档栏目管理模块 .....	44
5.6.1	文档版块栏目设置 .....	44
5.6.2	课件栏目的设置 .....	46
5.7	本章小结 .....	47
<b>第六章</b>	<b>系统测试 .....</b>	<b>48</b>
6.1	测试环境 .....	48
6.2	测试方法 .....	48
6.3	测试内容 .....	48
6.4	测试结果 .....	50



6.5 本章小结 .....	50
第七章 总结与展望 .....	51
7.1 总结 .....	51
7.2 展望 .....	51
参考文献 .....	52
致 谢 .....	54

## Contents

<b>Chapter 1 Introduction .....</b>	<b>1</b>
1.1 Research Background .....	1
1.2 Research Status .....	1
1.3 Main Content of the Paper .....	2
1.4 Outline of the Paper .....	2
<b>Chapter 2 Related Technical Introduction .....</b>	<b>4</b>
2.1 J2EE Technology Platform Overview .....	4
2.2 SSH2 Framework Overview .....	4
2.3 Struts2 Overview .....	5
2.4 Spring Overview .....	6
2.5 Hibernate Overview .....	7
2.6 Lucene Overview .....	8
2.7 Ik Analyzer Overview .....	11
2.8 Apache Solr Overview .....	11
2.9 Nginx Overview .....	12
2.10 Mysql Overview .....	13
2.11 Summary .....	14
<b>Chapter 3 System Requirements Analysis .....</b>	<b>15</b>
3.1 System Business Requirements Analysis .....	15
3.2 System User Requirements Analysis .....	15
3.3 System Functional Requirements Analysis .....	17
3.4 System Non-functional Requirements Analysis .....	19
3.5 Summary .....	19
<b>Chapter 4 System Design .....</b>	<b>20</b>
4.1 System Architecture Design .....	20
4.2 System Functional Design .....	23

4.3	System Database Design .....	23
4.4	Summary .....	27
<b>Chapter 5</b>	<b>System Implementation .....</b>	<b>28</b>
5.1	System Development And Runtime Environment .....	28
5.1.1	System Development Environment .....	28
5.1.2	System Runtime Environment .....	28
5.2	Document Conversion .....	28
5.3	Document Viewer .....	30
5.4	Document Retrieval .....	32
5.4.1	The Retrieval by Document Property .....	32
5.4.2	Theory of Full-text Search .....	32
5.4.3	Configuration of Full-text Retrieval Engine .....	35
5.4.4	Implementation of Full-text Retrieval .....	36
5.5	User Authority Management .....	38
5.5.1	User Login Authentication .....	38
5.5.2	User Password Reset.....	40
5.5.3	Role Management .....	41
5.5.4	User Management .....	41
5.6	Document Column Management .....	44
5.6.1	Configuration of Documentation Columns .....	44
5.6.2	Configuration of Courseware Columns .....	46
5.7	Summary .....	47
<b>Chapter 6</b>	<b>System Testing .....</b>	<b>48</b>
6.1	Test Environment .....	48
6.2	Test Methods .....	48
6.3	Test Contents .....	48
6.4	Test Results .....	50
6.5	Summary .....	50

<b>Chapter 7 Conclusions and Prospects .....</b>	<b>51</b>
<b>7.1 Conclusions .....</b>	<b>51</b>
<b>7.2 Future Task Prospects .....</b>	<b>51</b>
<b>References .....</b>	<b>52</b>
<b>Acknowledgements .....</b>	<b>54</b>

厦门大学博士论文摘要库

## 第一章 引言

### 1.1 研究背景和意义

随着网络科技的飞速发展,各种新的技术和应用不断出现,人们用各种方式通过网络来使得人们的工作学习和生活更加便利。人们可以在家同远在万里之外的亲友见面聊天,足不出户也可以购物甚至就餐,人们甚至可以轻而易举的发表自己的观点让所有人看见。网络技术的发展网络媒体的兴盛让信息的流动速度超乎想象。而巨量信息的存储及处理问题就变得愈发重要。在教学信息化中也存在着类似的问题。

教学方式信息化的过程中,老师上课的课件教案由手写的纸质材料变成了电子版的 PPT、WORD、PDF,学生的作业、毕业论文等很多也是以电子版的形式提交。学生每年毕业了又有新的学生进来,老师们的教学资料不断的修改更新,学院的 FTP 一次次的扩容,如何管理和维护这些数据资料,如何更有效的使用这些教学资料,让资料更方便有效的共享呢?这需要建立一个新的信息存储和检索系统,对目前已有的资料进行管理,也为以后的资料更新增加提供一个更有效的管理平台。这里核心需要实现的技术就是对于这些电子文档的全文检索功能。

### 1.2 国内外研究现状

根据信息的存储组织方式不同,可分为结构化数据和非结构化数据。对于结构化数据的存储和检索,通过数据库系统已经能够很好的来实现了。但是非结构化数据的存储管理及检索功能则很难直接通过数据库系统来完成。各种全文检索技术则是针对这一需求而产生的。全文检索技术也是目前网络检索的主要发展方向之一。一些成熟的技术框架也已出现,而现在使用最为广泛是 Apache 软件基金会 Jakarta 项目组的一个子项目 Lucene。它是一个开放源代码的全文搜索引擎工具包。Lucene 的目的是为软件开发人员提供一个简单易用的工具包,以方便地在目标系统中实现全文检索的功能,或者是以此为基础建立起完整的全文搜索引擎<sup>[1]</sup>。

当前已有不少基于 Lucene 开发的全文检索系统，比如毕业论文库全文检索系统<sup>[2]</sup>，科研文档全文检索系统<sup>[3]</sup>等。它们基本都实现了基本的全文检索功能，但主要只针对其系统中的毕业论文或科研文档进行检索，功能比较单一，缺乏相关的管理应用功能，对于检索得到的文档只提供了下载。

所以为了适应学院目前的实际需求情况，针对以上存在的问题，有必要研发一套新的具有全文检索、文档分类管理及在线阅读等功能的教学资源管理系统。本系统重点实现教学课件及相关文档的分类管理及全文检索、在线阅读功能，方便学院师生查阅相关资料。

### 1.3 论文的主要研究内容

本文要在阅读当前已有的全文检索系统相关介绍资料的基础上，针对学院对课件及毕业论文等教学和技术文档的存储、检索、查阅的需求，基于 J2EE 三层架构和 MySQL 数据库利用 Lucene 全文检索开发包设计并实现一套基于 J2EE 三层架构的具有可扩展性和伸缩性的教学资源管理系统。系统需要涵盖文档上传转换、文档查看、全文检索、系统管理等主要模块，重点解决目前各种教学文档资源检索和使用效率低下的问题，为师生的教学资源共享提供一个基础平台。

本文以软件工程理论为设计主线，需要说明教学资源管理系统的业务需求、功能需求、非功能需求、系统架构设计、功能设计和数据库设计等。并给出系统关键功能模块的代码实现过程以及系统的实现效果。

### 1.4 论文结构安排

本文共分为七章：

第一章，引言。首先介绍了本文的研究背景，随后分析了国内外的研究现状和发展趋势，最后是本文的组织结构。

第二章，相关技术介绍。对该教学资源管理系统所涉及的相关技术及概念进行阐述，并结合本系统的开发进行比较和分析。

第三章，系统需求分析。对教学资源管理系统进行需求分析，包括系统的业务需求、功能需求及非功能需求等。

第四章，系统总体设计。在系统需求分析的基础上进行功能模块划分，采用软件工程方法完成系统建模。

第五章，系统详细设计与实现。对系统设计进行细化分解，并介绍了系统的平台功能，着重介绍重点模块的实现方法

第六章，系统测试。介绍了常见的软件测试方法，并给出测试结果

第七章，总结与展望。总结工作和成果，指出不足，对系统进一步开发进行展望。

## 第二章 相关技术介绍

本系统是基于 J2EE 技术开发的，其中用到以 Struts2、Spring 及 Hibernate 为核心的 SSH2 框架。在全文搜索的功能实现方面采用了目前使用较多的一个开源引擎 Lucene，并加载了 IK Analyzer 工具包来支持中文的全文检索。系统的服务器架构上在前端使用了 Nginx 服务器提供高性能的静态网页资源访问，并将动态资源访问转发至 Tomcat 服务器。

### 2.1 J2EE 技术平台

J2EE 是企业级应用的一种规范，包含有很多方面的技术和构件：如 Javascript、Ajax、Jsp、Servlet、JavaBean、EJB、JMS、Jdbc、JTA、SOA、远程调用 RMI、事务、安全、容器、中间件、数据库以及开源的组件和框架，如 Ajax 框架 JQuery、展现层框架 Struts、业务逻辑框架 Spring、对象关系映像框架 Hibernate、Ibatis 等，SOA 规范的重重量级 Websevice 组件如 Xfire 组件、轻量级如 Hessian 组件，另外还有一些类似 JMS 和消息驱动 Bean 的异步消息处理框架等。另外还可以包含如异常错误及日志处理 Log4j 组件及 Applet 等客户端组件。<sup>[4]</sup>

### 2.2 SSH2 框架

SSH2 为 Struts2+Spring+Hibernate 的一个集成框架，是目前较流行的一种 Web 应用程序开源框架。集成 SSH2 框架的系统从职责上分为四层：表示层、业务逻辑层、数据持久层和域模块层，以帮助开发人员在短期内搭建结构清晰、可复用性好、维护方便的 Web 应用程序。其中使用 Struts2 作为系统的整体基础架构，负责 MVC 的分离，在 Struts2 框架的模型部分，控制业务跳转，利用 Hibernate 框架对持久层提供支持，Spring 做管理，管理 Struts2 和 hibernate。具体做法是：用面向对象的分析方法根据需求提出一些模型，将这些模型实现为基本的 Java 对象，然后编写基本的 DAO(Data Access Objects)接口，并给出 Hibernate 的 DAO



Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to [etd@xmu.edu.cn](mailto:etd@xmu.edu.cn) for delivery details.

厦门大学博硕士论文摘要库